

GridPP
UK Computing for Particle Physics



THE UNIVERSITY
of EDINBURGH

Edinburgh Site Report

Andrew Washbrook

University of Edinburgh

UK HEP SYSMAN Meeting

2nd June 2015

Support Structure

- The School of Physics and Astronomy Computing Services Team (CST) provide the majority of common computing support coverage for groups in the school:
 - Institute for Astronomy (IfA)
 - Institute for Condensed Matter and Complex Systems (ICMCS)
 - Institute for Particle and Nuclear Physics (IPNP)
 - Nuclear Physics
 - **Particle Physics Experiment**
 - Particle Physics Theory
 - EPCC
- Provide a choice of supported platforms across desktops, laptops, servers and Virtual Machines, to staff, students, visitors and projects
- By default, base supported platforms on University-wide services with local customisations to reflect the needs of the school

User Support

- Walk in helpdesk during normal working hours
 - Ticketing system through email and Unidesk Self service portal
-

System Deployment and Monitoring

- OS build deployment and configuration managed by LCFG
 - Exclusively Scientific Linux
- Manage over **300** desktops and **150** servers (including VMs) across the school

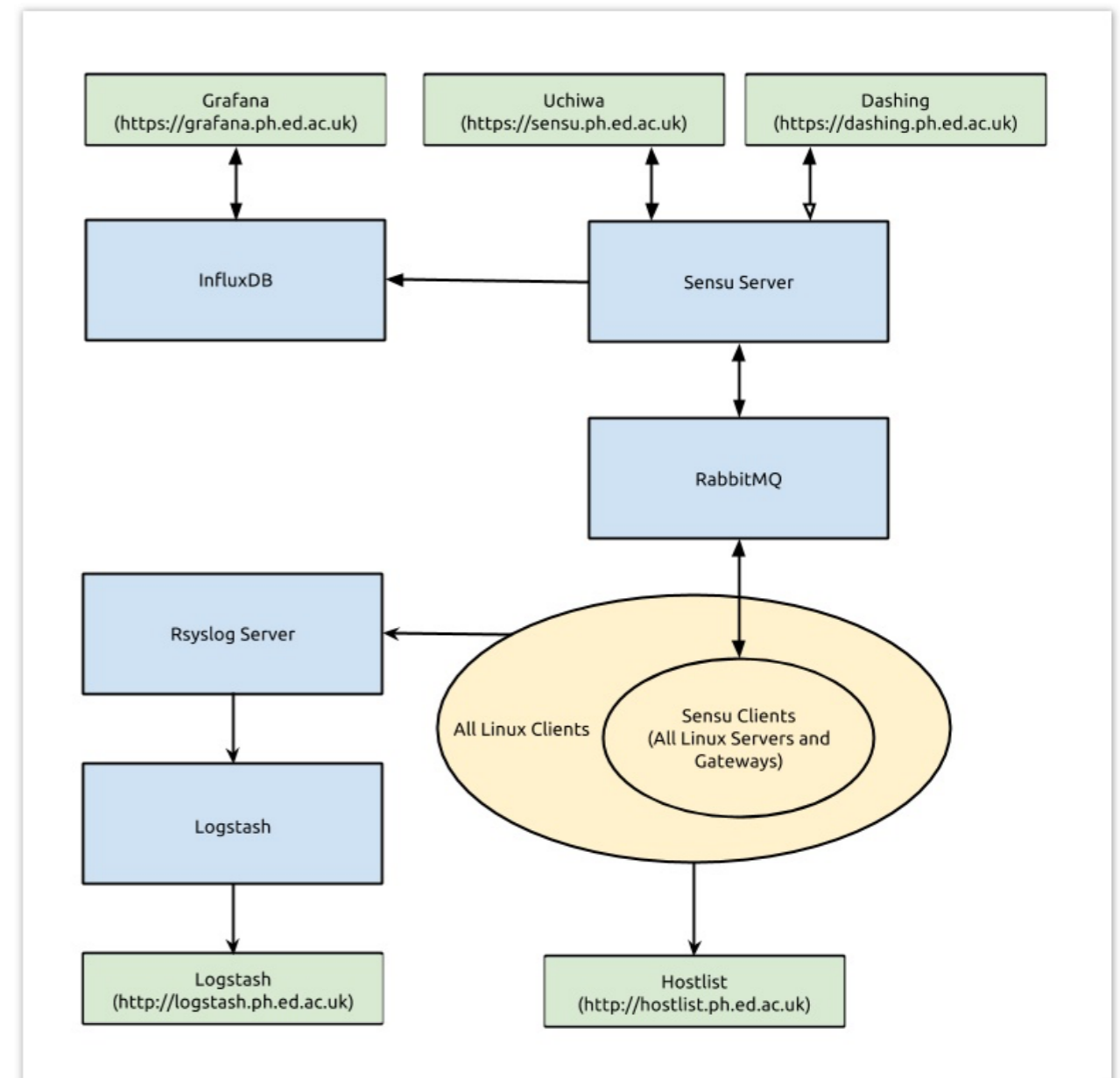
System Monitoring

- **Sensu, Dashing, Uchiwa** and **Grafana** are used to monitor system activities
- Logstash used to capture syslog, auditd and fail2ban messages across the estate

SL7 Build

- SL7 build already deployed as part of an early adoption program
- Currently evaluating the build for Grid and LHC experiment software
- No timeframe set for full-scale deployment - some services will be migrated over the summer

System Monitoring Flow



System Monitoring Examples

uchima

EVENTS >

ALL DATACENTERS ACTIONS HIDE 10 OF 10

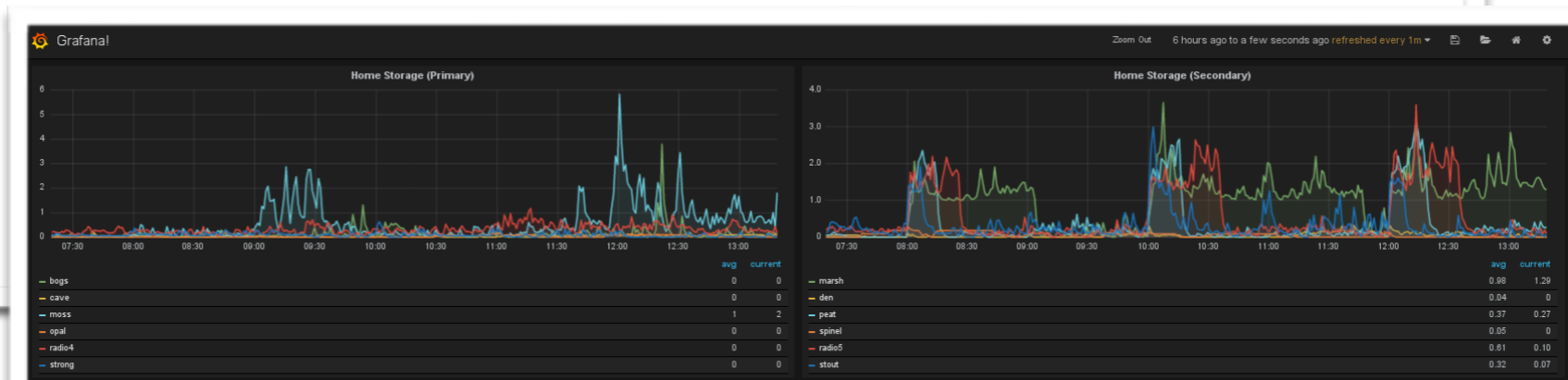
Search

Source	Check	Output	Count	Alert	Time
ngc02.ph.ed.ac.uk	ceph	CheckCephHealth CRITICAL: Error initializing cluster client: Error(error calli...	231660	Sensu	a few seconds ago
pyke.ph.ed.ac.uk	disk	CheckDisk CRITICAL: /scratch/c 100%	20445	Sensu	a few seconds ago
botley.ph.ed.ac.uk	mysql	CheckMySQL CRITICAL: Error message: Access denied for user 'root'@'loc...	14377	Sensu	a minute ago

Source:

- ngc02.ph.ed.ac.uk
- pyke.ph.ed.ac.uk
- botley.ph.ed.ac.uk
- vine.ph.ed.ac.uk
- morar2.ph.ed.ac.uk
- auntie.ph.ed.ac.uk
- winkleman.ph.ed.ac.uk
- mormont.ph.ed.ac.uk
- feltz.ph.ed.ac.uk
- statix.ph.ed.ac.uk

Uchiwa



Grafana

kibana Discover Visualize Dashboard Settings

Syslog

Syslog - Programs

Syslog - Bar

Syslog

Count

kernel dhttpd CFKOND sethd apwatch sendmail automount borouffe gnome-session sudo

Top 10 syslog_program.raw

Syslog - Severity

Count

debug informational warning error notice

Top 5 syslog_severity.raw

June 1st 2015, 13:20:18.000 gaunt kernel ipt IN=eth0 OUT= MAC=ff:ff:ff:ff:a0:48:1c:8f:01:71:08:00 SRC=129.215.73.106 DST=255.255.255.255 LEN=68 TOS=0x00 PREC=0x00 TTL=128 ID=2969 PROTO=UDP SPT=57250 DPT=1947 LEN=48

June 1st 2015, 13:20:18.000 router kernel DROP_inbound-IntCatchAll IN=eth6 OUT=eth7 SRC=124.232.142.220 DST=129.215.73.213 LEN=60 TOS=0x00 PREC=0x00 TTL=238 ID=54321 PROTO=UDP SPT=53365 DPT=53 LEN=40

June 1st 2015, 13:20:18.000 grell kernel ipt IN=eth0 OUT= MAC=ff:ff:ff:ff:a0:48:1c:8f:01:71:08:00 SRC=129.215.73.106 DST=255.255.255.255 LEN=68 TOS=0x00 PREC=0x00 TTL=128 ID=2969 PROTO=UDP SPT=57250 DPT=1947 LEN=48

June 1st 2015, 13:20:18.000 router kernel DROP_inbound-IntCatchAll IN=eth6 OUT=eth7 SRC=115.51.232.114 DST=129.215.75.211 LEN=69 TOS=0x00 PREC=0x00 TTL=47 ID=0 DF PROTO=UDP SPT=40420 DPT=9909 LEN=49

June 1st 2015, 13:20:17.000 router kernel DROP_inbound-IntCatchAll IN=eth6 OUT=eth7 SRC=80.85.84.75 DST=129.215.75.233 LEN=40 TOS=0x00 PREC=0x00 TTL=244 ID=57456 PROTO=TCP SPT=40001 DPT=5900 WINDOW=1024 RES=0x00 SYN URGP=0

Kibana 4

PPE Group Perspective

- I am the "principle user" for the PPE group to communicate group requirements and concerns
- Support issues mostly common to other groups but there are PPE-specific requirements:
 - *Essential Tools and Packages*: ROOT, VidyO, domain specific compilers (icc, clang)
 - *Filesystems*: CVMFS + Squid proxy, AFS
 - *Connectivity and Access*:
 - High redundancy PPE gateway service for remote users
 - Access to CERN services (ssh, kerberos, AI cloud services, Cernbox)
 - Job submission and data retrieval from the LHC computing grid
 - Selected access to group-specific hardware (labs and development testbeds)

User Community

- Traditionally supported Tier-2 and Tier-3 resources for ATLAS and LHCb
 - Equates to roughly 50/50 split of researchers in the group
 - Now include new group interests: Hyper-K and Future colliders (CLIC, FCC, ILC)
 - Increasing engagement with other projects and experiments in the School:
 - Lux-Zeplin experiment (LZ)
 - Large Synoptic Survey Telescope (LSST)
 - EUCLID Telescope
-

Eddie ECDF Cluster

- Use ECDF facility for both a Tier-2 Grid computing and as a Tier-3 group resource

Eddie Mark 2

- Phase 1 - 130 x IBM dx360M3 iDataPlex servers (2 x Xeon E5620 quad-core)
- Phase 2 - 156 x IBM dx360M3 iDataPlex servers (2 x Xeon E5645 six-core)
 - ~3,000 cores
- GPU and large memory systems
- Single queue for single core and multi-core workloads
- Equipment hosting provided by Advanced Computing Facility (ACF)
- ACF provide:
 - Infrastructure management
 - Power
 - Cooling
 - Security
 - Routine system tasks (e.g. disk replacement)



ECDF Customer Base

- ECDF provides computing resources across the university for: **Physics, Geoscience, Engineering, Life sciences, Veterinary medicine, Informatics and Biology**

Support and feedback

- Ticketing system manages incident response calls and simple change requests
- Ongoing collaboration with groups that have bespoke requirements (i.e. GridPP)
- Software troubleshooting and optimisation
- Regular drop-in sessions

Benefits to PPE Group

- We don't have to care about:
 - Cluster and batch system setup and configuration
 - Continual equipment maintenance
 - System wide troubleshooting
 - Occasional leveraging of opportunistic resources
 - Better resource size to FTE ratio
-

Eddie Mark 3

Eddie Mark 3 - Available from August 2015

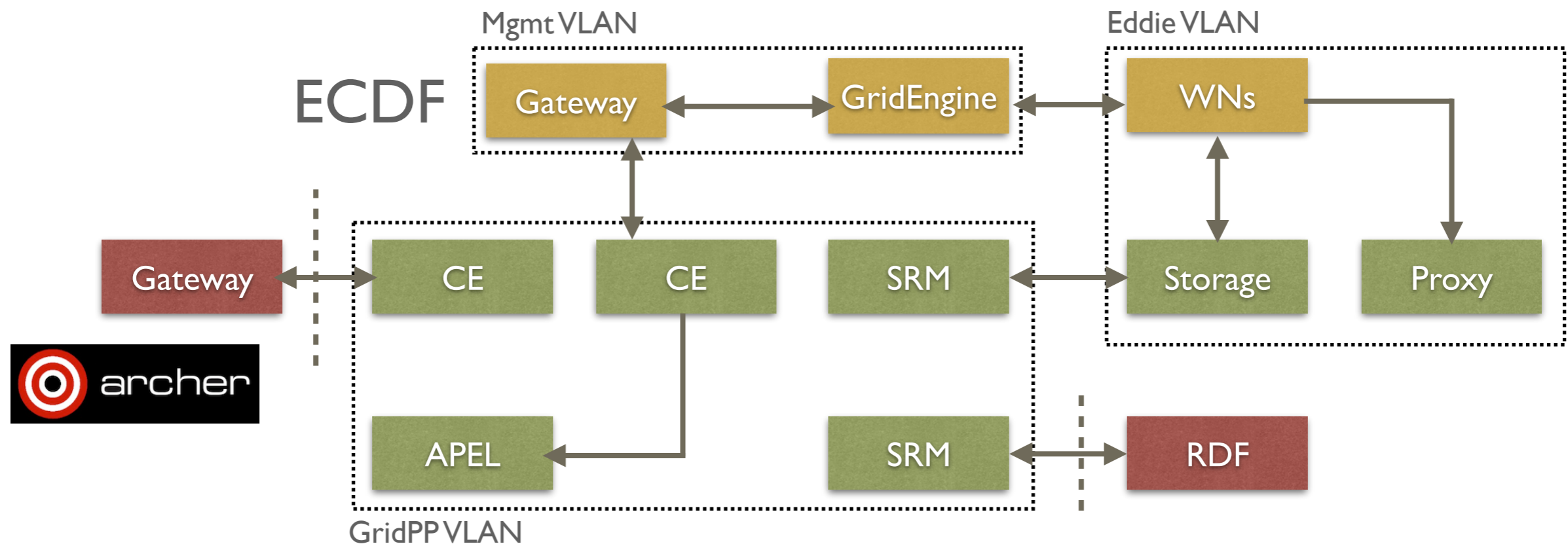
- Tendered for £1M of new equipment
 - We had early input into machine specification for PPE
- Similar operational model
 - "Free" at point of use
 - Paid-for jobs have higher priority
 - Opportunistic use encouraged
- Hosting service for additional compute purchased by university research groups
 - Spare rack capacity for bespoke equipment

Details (New)

- **368** compute nodes with 2 x 2 TB large memory nodes
 - 113 exclusively for Institute of Genetics and Molecular Medicine (IGMM))
 - Each node has 2 x **E5-2630v3** cpus (8 cores per chip = total **5,888** cores)
 - 2 nodes with 2 x Nvidia Tesla K80 GPUs
 - Mostly 4 GB memory per core, select machines have higher ratio
 - All connected via a 10/40 GE network.
 - Two storage platforms: **300 TB** (11 GB/s) and **633 TB** (7GB/s)
-

Archer and RDF

- The PPE group has access to the Archer supercomputer as part of University directors time
- Ongoing work to integrate access to Tier-2 middleware to enable a persistent distributed computing resource for LHC experiments
- Storage from UK Research Data Facility is integrated in Tier-2 operations
- Archer also used for running short-term standalone mini-simulation challenges
 - Large scale simulation for ATLAS exotic physics studies is underway



Archer and RDF Tier-2 Integration

Other Computing Resources

Tier-3 School cluster

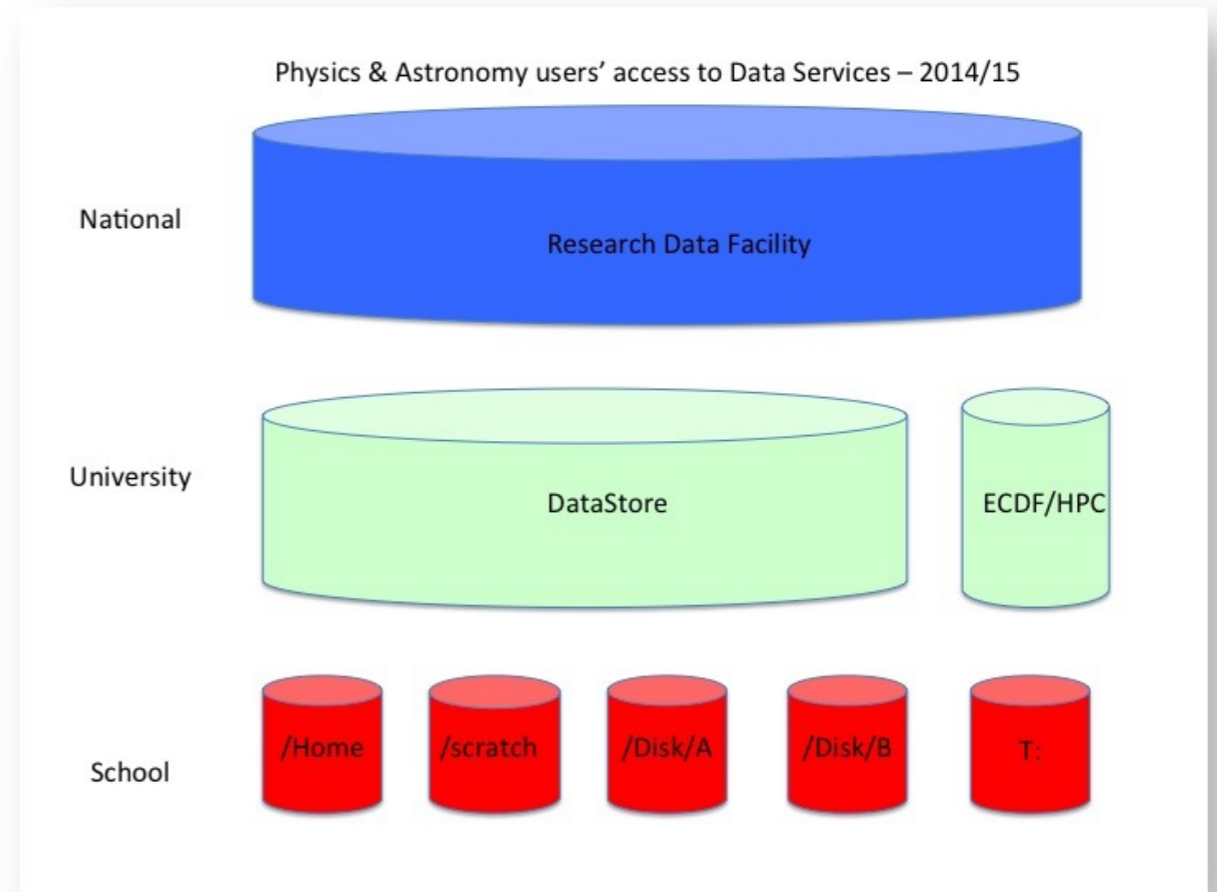
- In the process of purchasing a cluster to consolidate small scale computing needs across all the physics groups in the school
- Aim to augment access to central university resources (such as ECDF)
- More or less similar requirements apart from memory per core

Development Testbed

- Private testbed available bespoke PPE studies:
 - ATLAS High Level Trigger profiling and optimisation studies
 - Many-core platform development for Trigger algorithms and the acceleration of supervised learning techniques
 - Multi-core workload scheduling studies
 - 4 Dell servers + Supermicro
 - Nvidia K40 GPU
 - Intel Xeon Phi
-

Tier-3 Storage

- By default, individuals can access two layers of Data services:
 - School storage
 - University storage (**DataStore**).
- The School Storage layer provides small-scale, backed-up local storage, usually organised per research group, and often coupled to local compute resources



PPE Group Requirements

- School storage primarily for home directories and legacy scratch space
- Currently using **35 TB** of additional space for dataset caching and archiving (recurring charge per TB/year)
- Exploring use of GfalFS on desktop clients for direct access to our **1 PB** of local Grid storage

University DataStore Scheme

- University introduced **DataStore** scheme in 2014 as a free at point of use allocation of **0.5 TB** per researcher as a general purpose file service.
- Each researcher (at their individual discretion) can choose to share up to **50%** of their individual allocation
- Complements the University's **DataShare** service, which enables researchers to publish research data, create and manage identifiers and citation references

Details

- **3.8 PB** total "usable" disk capacity, plus around **5 PB** tape (for backup)
- GPFS, with TSM for backups
- Daily ~ **7,000-11,000** active users
- Working to closer integration with ECDF
- Backup:
 - 10 days' online file history
 - 60 days' offline backup history
 - Full offsite replica for disaster recovery



DataSync (New)

- Dropbox-like file hosting service for non-sensitive data
- 5 GB storage plus connection to personal and group data share

Any Questions?
